



City Research Online

City, University of London Institutional Repository

Citation: MacFarlane, A. (2013). Web Search Engine Research. Journal of Documentation, 69(4), pp. 594-596. doi: 10.1108/JD-04-2013-0047

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/4505/>

Link to published version: <http://dx.doi.org/10.1108/JD-04-2013-0047>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Book Review

Title: Web Search Engine Research

Author: Dirk Lewansowski (ed)

Reviewer: A. MacFarlane, City University London

Review:

The web has been around for nearly 20 years now and an important enabling technology is web search. Change is constant in the web, and it is essential to continue to innovate in it, understanding its impact and the disruption the technology causes. This book gives a broad overview of three areas of web search namely: emerging areas, thinking beyond traditional evaluation methods and new perspectives. The purpose of the book is to provide an overview of issues in each of these key areas pointing to further research as and when necessary. It is an edited book, which brings in contributions from experts in the three stated themes which is a key strength in my view. The introduction by Lewandowski provides a useful overview of the book and its themes.

Three emerging areas are tackled in the book: social search, local search and web search statistical analysis. Burghardt et al provide an analysis of social search, its many definitions and history/background etc. There are many different forms of social search e.g. people search, search using social data etc. They discuss the ambiguity of the phrase 'social search' – the definition largely depends on the application. Examples of real social search systems are provided. Related issues such as a question-answering and collaborate/personalised search/filtering are addressed. I found this chapter particularly useful. Ahlers addresses the issue of local search i.e. the use of geographical information retrieval techniques for spatial-temporal information. The focus is very much on maps, and the display of results, with a discussion on how such technologies can be combined in web search. Mobile search is briefly tackled. Pietruszkiewicz focuses on the analysis of web search statistics i.e. analysis of searches already carried out – trends, relations, economic impact etc. How is this data collected and analysed? A very good example of an application is Google trends (<http://www.google.com/trends/>), which allows a user to examine the impact of a search term or phrase over time and space.

Four issues in terms of methodological issues in evaluation are tackled: evaluating effectiveness, the impact of diversity, personalised search and rank correlation. Carterette et al review the concept of test collection and metrics – much studied in the field for around 40 years – and look to more user focused evaluation methods and metrics e.g. examining logs of online sessions. Deneck reviews the concept of diversity. Many web search engines attempt to resolve a query by returning one interpretation of a given topic e.g. virus – which can be a medical or computing concept. In some cases it is desirable to retrieve documents with multiple interpretations or different aspects of the same search term. A review of diversity in web search is provided and an example put forward for the provision of diverse results in the medical domain. Fun et al review personalised search (Google does this for most searches now) and they provide a model for evaluating and comparing this type of search, with a focus on the metrics and relevance assessment issues. Melucci focuses on the search engine results page (SERP) in terms of whether or not two search hitlists are similar – statistical rank correlation methods can be used to examine this issue.

Three new perspectives on web search are examined: looking beyond search, a user perspective on web search evaluation and examining more theoretical aspects of web search in terms of epistemology etc. Bryant et al review the traditional use of search engines and looks beyond this by examining probes – either viewing results as cards or pebbles over the retrieved information. The focus therefore is very much on new HCI methods for web search. Kammerer and Gerjets focuses on how end users actually evaluation web search engine results and their behaviour in assessing the value of what is returned. This information is used to make suggestions as to how the credibility of user assessments of SERPS can be improved. Rall takes a very philosophical approach to search, using epistemology and concepts in philosophy of science to examine what is actually meant by the term 'search'.

Who would find this book useful? I think it is a valuable resource for both established researchers and students, and I would recommend the book to both groups. Researchers in web search will find the book a key text for their work, whilst it provides students with a good background in many aspects of the field – for both computer science and library and information science. It is clearly a useful text for information retrieval modules, and I will be using various chapters for the module that I deliver (social search is an example). The book is grounded in a firm body of knowledge in the field, built up over many years – but does use this as a base to examine new and upcoming ideas both in terms of technology and evaluation of that technology. I recommend it highly.